Docket No.: 08204/1200311-US2
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Letters Patent of:
Paul I. Szabo et al.

Patent No.: 7,395,349

Issued: July 1, 2008

For: METHOD AND SYSTEM FOR SCALING NETWORK TRAFFIC MANAGERS

REQUEST FOR CERTIFICATE OF CORRECTION
PURSUANT TO 37 CFR 1.323 AND 1.322

Attention: Certificate of Correction Branch
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Upon reviewing the above-identified patent, Patentee noted typographical errors which should be corrected. A listing of the errors to be corrected is attached.

The typographical errors marked with an "A" on the attached list are found in the application as filed by applicant. Please charge our Credit Card in the amount of $100.00 covering the fee set forth in 37 CFR 1.20(a).

The typographical errors marked with a "P" on the attached list are not in the application as filed by applicant. Also given on the attached list are the documents from the file history of the subject patent where the correct data can be found.

The errors now sought to be corrected are inadvertent typographical errors the correction of which does not involve new matter or require reexamination.

Transmitted herewith is a proposed Certificate of Correction effecting such corrections. Patentee respectfully solicits the granting of the requested Certificate of Correction.

The Commissioner is authorized to charge any deficiency of up to $300.00 or credit any excess in this fee to Deposit Account No. 04-0100.

Dated: July 28, 2008

Respectfully submitted,

By _____

Flynn Barrison
Registration No.: 53,970
DARBY & DARBY P.C.
P.O. Box 770
Church Street Station
New York, New York 10008-0770
(212) 527-7700
(212) 527-7701 (Fax)
Attorneys/Agents For Applicant

PTO/SB/44 (09-07)
Approved for use through 08/31/2010. OMB 0651-0033
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

(Also Form PTO-1050)

UNITED STATES PATENT AND TRADEMARK OFFICE
CERTIFICATE OF CORRECTION

PATENT NO. : 7,395,349 Page 1 of 1

APPLICATION NO.: 10/644,692

ISSUE DATE : 10/644,692

INVENTOR(S) : Szabo et al.

It is certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

On the face page, in field (56), under "Other Publications", in column 2, line 4, delete "1-173" and insert - - 1-173. - -, therefor.

On the face page, in field (56), under "Other Publications", in column 2, line 7, delete "Hoops" and insert - - Hopps - -, therefor.

In column 1, line 4, after "is" insert - - a - -.

In column 3, lines 15-20, delete "Furthermore, computers, such as remote computer 140, and other related electronic devices can be remotely connected to either LANs 120.sub.a-d or WAN 130 via a modem and temporary telephone link. The number of WANs, LANs, and routers in FIG. 1 may be increased or decreased arbitrarily without departing from the spirit or scope of this invention." and insert the same at Line 14 as a continuation of the same paragraph.

In column 6, lines 6-7, delete "NETheui" and insert - - NETbeui - -, therefor.

In column 6, line 61, delete "requesters" and insert - - requestors - -, therefor.

In column 7, line 39, delete "gov/and" and insert - - gov/ and - -, therefor.

In column 18, line 26, in Claim 26, after "using" delete "the" and insert - - in - -, therefor

MAILING ADDRESS OF SENDER (Please do not use customer number below):
John W. Branch, Esq.
DARBY & DARBY P.C.
P.O. Box 770
Church Street Station
New York, New York 10008-0770

1

DARBY & DARBY

Issued Patent Proofing Form File#: **08204/1200311-US2**

Note: **P** = USPTO Error **A** = Applicant Error

US Serial No.: **10/644,692** US Patent No.: **US 7,395,349 B1** Issue Date: **Jul. 1, 2008**

Title: **METHOD AND SYSTEM FOR SCALING NETWORK TRAFFIC MANAGERS**

S. No.	P/A	Original		Issued Patent		Description of Error
		Page	Line	Column	Line	
1	P	Sheet 1 of 1 List of References cited by applicant and considered by examiner (10/30/2007)	Entry 1 Line 3 (Non Patent Literature Documents)	First Page Col. 2 (Other Publications)	4	Delete "1-173" and insert - - 1-173. - -, therefor.
2	P	Sheet 1 of 1 List of References cited by applicant and considered by examiner (10/30/2007)	Entry 3 Line 1 (Non Patent Literature Documents)	First Page Col. 2 (Other Publications)	7	Delete "Hoops" and insert - - Hopps - -, therefor.
3	P	Page 2 Amendment After Final or under 37CFR 1.312, initialed by the examiner. (02/26/2008)	3 (Amendments to the Specification)	1	4	After "is" insert - - a - -.
4	P	Page 4 Specification (08/20/2003)	11-15	3	15-20	Delete "Furthermore, computers, such as remote computer 140, and other related electronic devices can be remotely connected to either LANs 120.sub.a-d or WAN 130 via a modem and temporary telephone link. The number of WANs, LANs, and routers in FIG. 1 may be increased or decreased arbitrarily without departing from the spirit or scope of this invention." and insert the same at Line 14 as a continuation of the same paragraph.
5	P	Page 9 Specification (08/20/2003)	1	6	6-7	Delete "NETheui" and insert - - NETbeui - -, therefor.

6	P	Page 10 Specification (08/20/2003)	12	6	61	Delete "requesters" and insert - - requestors - -, therefor.
7	P	Page 11 Specification (08/20/2003)	14	7	39	Delete "gov/and" and insert - - gov/ and - -, therefor. (Consider Space)
8	A	Page 6 Claims (01/29/2008)	Claim 27 Line 1	18	26	In Claim 26, after "using" delete "the" and insert - - in - -, therefor

US007395349B1

(12) **United States Patent**
Szabo et al.

(10) **Patent No.: US 7,395,349 B1**
(45) **Date of Patent: *Jul. 1, 2008**

(54) **METHOD AND SYSTEM FOR SCALING NETWORK TRAFFIC MANAGERS**

(75) Inventors: **Paul Szabo**, Seattle, WA (US); **David D. Schmitt**, Seattle, WA (US); **Ning X. Li**, Sequim, WA (US)

(73) Assignee: **F5 Networks, Inc.**, Seattle, WA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 856 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **10/644,692**

(22) Filed: **Aug. 20, 2003**

Related U.S. Application Data

(63) Continuation-in-part of application No. 10/119,433, filed on Apr. 9, 2002, now Pat. No. 7,102,996.

(60) Provisional application No. 60/293,466, filed on May 24, 2001.

(51) **Int. Cl.**
G06F 15/173 (2006.01)

(52) **U.S. Cl.** **709/238**

(58) **Field of Classification Search** 709/238, 709/245

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

6,157,950	A *	12/2000	Krishnan	709/223
6,182,146	B1 *	1/2001	Graham-Cumming, Jr.	709/238
6,742,045	B1 *	5/2004	Albert et al.	709/238
6,775,235	B2 *	8/2004	Datta et al.	370/238
2003/0229809	A1 *	12/2003	Wexler et al.	713/201
2005/0008017	A1 *	1/2005	Datta et al.	370/392

OTHER PUBLICATIONS

"Amendment to Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications—Aggregation of Multiple Link Segments", IEEE Std. 802.3ad-2000, Mar. 30, 2000, pp 1-173

Thaler, D., et al., "Multiple Issues in Unicast and Multicast Next-Hop Selection", The Internet Society, Nov. 2000, pp. 1-9.

Hoops C., "Anaylsis of an Equal-Cost Multi-Path Algorithm", The Internet Society, Nov. 2000, pp. 1-8.

Moy, J., "OSPF Version 2", The Internet Society, Apr. 1998, pp. 1-53.

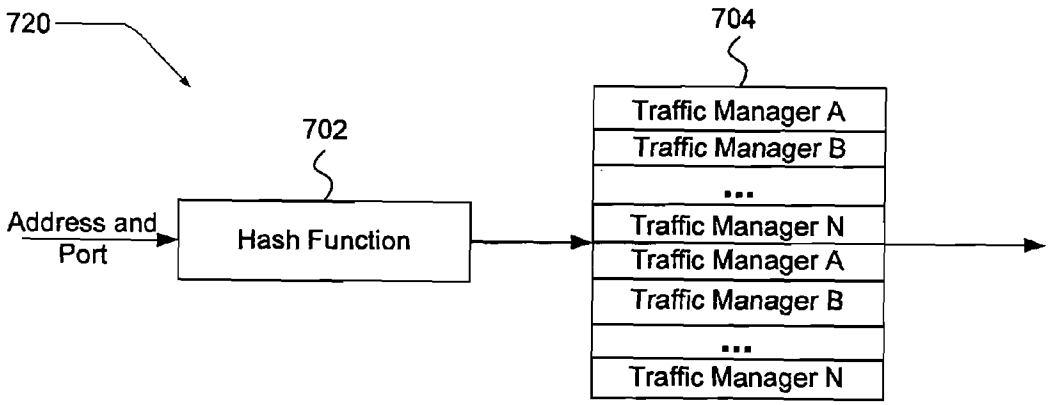
* cited by examiner

Primary Examiner—Salad E Abdullahi
(74) *Attorney, Agent, or Firm*—Darby & Darby P.C.; Jamie L. Wiegand

(57) **ABSTRACT**

A method and system is directed to routing a flow of packets over a network to multiple traffic management devices. An apparatus receives each packet from a network and forwards the packet to one of a group of traffic management devices. The apparatus also may receive packets from servers for which the traffic management devices are managing communications. When forwarding packets, a traffic management device is selected from the group of traffic management devices by employing a hash of an IP address and port number. The IP address and port number are selected from source or destination information in the packet that has a greater port number. When the traffic management device performs a network address translation, further actions may be performed so that packets that are part of a flow between two network devices are delivered to the same traffic management device.

29 Claims, 9 Drawing Sheets



METHOD AND SYSTEM FOR SCALING NETWORK TRAFFIC MANAGERS

This application is a continuation-in-part of U.S. patent application Ser. No. 10/119,433 now U.S. Pat. No. 7,102,996, filed Apr. 9, 2002, which claims the benefit of priority under 35 U.S.C § 119(e) from U.S. Provisional Application Ser. No. 60/293,466, filed May 24, 2001, entitled "Method and System for Distributing Traffic to Multiple Load Balancers," the benefit of each is further claimed herein.

FIELD OF THE INVENTION

The present invention relates to computer network traffic, and in particular to distributing network traffic associated with traffic management devices.

BACKGROUND

The Internet's core bandwidth continues to double every year. Some of this additional bandwidth is consumed as more and more users access the Internet. Other additional bandwidth is consumed as existing users increase their use of the Internet. This increase of Internet use translates into an increase in traffic directed to and from World Wide Web (WWW) servers and other Internet servers.

Replacing a WWW server with a WWW server of twice the capacity is a costly undertaking. Adding additional WWW servers is less costly but generally requires a load-balancing mechanism to balance workload so that each virtual server performs work proportional to its capacity and the number of servers available to the traffic management device that is performing the load balancing.

This requirement for more sophisticated traffic management requires more processing. With a sufficient rate of requests, eventually a traffic management device may not be able to process traffic in a timely manner. Therefore, it is with respect to these considerations and others that the present invention has been made.

BRIEF DESCRIPTION OF THE DRAWINGS

Non-limiting and non-exhaustive embodiments of the present invention are described with reference to the following drawings. In the drawings, like reference numerals refer to like parts throughout the various figures unless otherwise specified.

For a better understanding of the present invention, reference will be made to the following Detailed Description of the Invention, which is to be read in association with the accompanying drawings, wherein:

FIGS. **1-2** show components of an exemplary environment in which the invention may be practiced;

FIG. **3** illustrates an exemplary environment in which a system for distributing traffic to an array of traffic management devices operates;

FIG. **4** shows another exemplary environment in which a system for distributing traffic to an array of traffic management devices operates;

FIG. **5** shows yet another exemplary environment in which a system for routing traffic through an array of traffic management devices operates;

FIG. **6** shows one embodiment of a statistical traffic distributor (STD);

FIG. **7** shows another embodiment of a statistical traffic distributor (STD);

FIG. **8** illustrates a flow chart for a process for determining how to forward packets; and

FIG. **9** shows a flow chart for managing target port number selection, in accordance with the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

The present invention now will be described more fully hereinafter with reference to the accompanying drawings, which form a part hereof, and which show, by way of illustration, specific exemplary embodiments by which the invention may be practiced. This invention may, however, be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. Among other things, the present invention may be embodied as methods or devices. Accordingly, the present invention may take the form of an entirely hardware embodiment, an entirely software embodiment or an embodiment combining software and hardware aspects. The following detailed description is, therefore, not to be taken in a limiting sense.

Throughout the specification, the meaning of "a," "an," and "the" include plural references. The meaning of "in" includes "in" and "on."

Briefly stated, the present invention is directed to a system and method for routing a flow of packets to one or more traffic management devices. The invention enables a client-side flow of packets and a server-side flow of packets in the flow of packets between a client and server to be directed to a same traffic management device. One or more distributors may be employed to route the flow of packets. As the packets are received, a distributor extracts information, such as source and destination IP addresses, and source and destination port numbers from the packet. If the source port number equals the destination port number, the distributor forwards the packet to a pre-determined traffic management device. If the source port number is greater than the destination port number, the distributor performs a hash on the source IP address and the source port number. If the source port number is not greater than the destination port number, the distributor performs the hash on the destination IP address and the destination port number. For both cases, the hash results are employed to select a destination traffic management device to forward the packet. Because, the destination traffic management device may perform a network address translation on the packet contents, the present invention may perform additional actions directed at maintaining the same relationship between the source and destination port numbers, such that a reply packet to the forwarded packet is handled by the same traffic management device.

Illustrative Operating Environment

FIGS. **1-2** show components of an exemplary environment in which the invention may be practiced. Not all of the components may be required to practice the invention, and variations in the arrangement and type of the components may be made without departing from the spirit or scope of the invention.

FIG. **1** shows one embodiment of wide area network/local area network (WAN/LAN) **100**, in accordance with the present invention. WAN/LAN **100** includes a plurality of local area networks ("LANs") $\mathbf{120}_{a-d}$ and wide area network ("WAN") **130** interconnected by routers **110**. Routers **110** are intermediary devices on a communications network that

expedite message delivery. On a single network linking many computers through a mesh of possible connections, a router receives transmitted messages and forwards them to their correct destinations over available routes. On an interconnected set of LANs—including those based on differing architectures and protocols—, a router acts as a link between LANs, enabling messages to be sent from one to another. Communication links within LANs typically include twisted pair, fiber optics, or coaxial cable, while communication links between networks may utilize analog telephone lines, full or fractional dedicated digital lines including T1, T2, T3, and T4, Integrated Services Digital Networks (ISDNs), Digital Subscriber Lines (DSLs), wireless links, or other communications links known to those skilled in the art.

Furthermore, computers, such as remote computer **140**, and other related electronic devices can be remotely connected to either LANs **120**$_{a-d}$ or WAN **130** via a modem and temporary telephone link. The number of WANs, LANs, and routers in FIG. **1** may be increased or decreased arbitrarily without departing from the spirit or scope of this invention.

As such, it will be appreciated that the Internet itself may be formed from a vast number of such interconnected networks, computers, and routers. Generally, the term "Internet" refers to the worldwide collection of networks, gateways, routers, and computers that use the Transmission Control Protocol/Internet Protocol ("TCP/IP") suite of protocols to communicate with one another. At the heart of the Internet is a backbone of high-speed data communication lines between major nodes or host computers, including thousands of commercial, government, educational, and other computer systems, that route data and messages. An embodiment of the invention may be practiced over the Internet without departing from the spirit or scope of the invention.

The media used to transmit information in communication links as described above illustrates one type of computer-readable media, namely communication media. Generally, computer-readable media includes any media that can be accessed by a computing device. Computer-readable media may include computer storage media, communication media, or any combination thereof.

Communication media typically embodies computer-readable instructions, data structures, program modules, or other data in a modulated data signal such as a carrier wave or other transport mechanism and includes any information delivery media. The term "modulated data signal" means a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, communication media includes wired media such as twisted pair, coaxial cable, fiber optics, wave guides, and other wired media and wireless media such as acoustic, RF, infrared, and other wireless media.

Remote computer **140** is any device capable of connecting with local area networks ("LANs") **120**$_{a-d}$ and wide area network ("WAN") **130**. The set of such devices may include devices that typically connect using a wired communications medium such as personal computers, multiprocessor systems, microprocessor-based or programmable consumer electronics, network PCs, and the like. The set of such devices may also include devices that typically connect using a wireless communications medium such as cell phones, smart phones, pagers, walkie talkies, radio frequency (RF) devices, infrared (IR) devices, CBs, integrated devices combining one or more of the preceding devices, and the like. Alternatively, remote computer **140** may be any device that is capable of connecting using a wired or wireless communication medium such as a PDA, POCKET PC, wearable computer, or other device mentioned above that is equipped to use a wired and/or wireless communication medium.

FIG. **2** shows an exemplary network device **200** that may operate as an intermediate network device in accordance with the present invention. It will be appreciated that not all components of network device **200** are illustrated, and that network device **200** may include more or fewer components than those shown in FIG. **2**. Network device **200** may operate, for example, as a router, bridge, firewall, gateway, traffic management device (also referred to as a traffic manager), distributor, load balancer, server array controller, or proxy server. The communications may take place over the network **130**, the Internet, a WAN, LAN, or some other communications network known to those skilled in the art.

As illustrated in FIG. **2**, network device **200** includes a central processing unit (CPU) **202**, mass memory, and a network interface unit **212** connected via a bus **204**. Network interface unit **212** includes the necessary circuitry for connecting network device **200** to network **130**, and is constructed for use with various communication protocols including the TCP/IP and UDP/IP protocol. Network interface unit **212** may include or interface with circuitry and components for transmitting messages and data over a wired and/or wireless communications medium. Network interface unit **212** is sometimes referred to as a transceiver.

The mass memory generally includes random access memory ("RAM") **206**, read-only memory ("ROM") **214**, and one or more permanent mass storage devices, such as hard disk drive **208**. The mass memory stores operating system **216** for controlling the operation of network device **200**. The operating system **216** may comprise an operating system such as UNIX, LINUX™, or Windows™.

In one embodiment, the mass memory stores program code and data for implementing a hash function **218**, and program code and data for implementing an allocation table **220**, in accordance with the present invention. The mass memory may also store additional program code **224** and data for performing the functions of network device **200**.

In one embodiment, the network device **200** includes one or more Application Specific Integrated Circuit (ASIC) chips **226** connected to the bus **204**. As shown in FIG. **2**, the network interface unit **212** may connect to the bus through an ASIC chip. The ASIC chip **226** includes logic that performs some of the functions of network device **200**. For example, in one embodiment, the ASIC chip **226** performs a number of packet processing functions, to process incoming packets. In one embodiment, the logic of the hash function **218** is performed by the ASIC chip **226**. In one embodiment, the network device **200** includes one or more field-programmable gate arrays (FPGA) (not shown), instead of, or in addition to, the ASIC chip **226**. A number of functions of the network device can be performed by the ASIC chip **226**, by an FPGA, by the CPU **202** with the logic of program code stored in mass memory, or by a combination of the ASIC chip and the CPU.

Computer storage media may include volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information, such as computer readable instructions, data structures, program modules or other data. Examples of computer storage media include RAM **206**, ROM **214**, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium that can store the information and that can be accessed by a computing device.

Network device **200** may also include an input/output interface (not shown) for communicating with external devices or users.

Network device **200** can also be implemented as one or more "blades" where the term "blade" refers to one of multiple electronic circuit boards or cards that are installed in a hardware chassis with a backplane. An exemplary blade may include one or more processors, volatile and non-volatile memory, interfaces suitable for communicating information to and from the blade, and other components for enabling the operation of one or more applications. A blade may also include a specialized interface for the backplane and other interfaces, such as a USB port, FIREWIRE port, serial port, RF interface, IR interface, Ethernet interface, IDE controller, and the like. An application running on a blade may employ any of these interfaces to communicate information to other applications running on other blades and/or devices coupled to the blade server. Network device **200** can also be implemented as a combination of blades and additional components in the chassis.

Illustrative Traffic Distributing Systems

FIG. **3** illustrates an exemplary environment in which a system for distributing traffic through an array of traffic management devices operates, according to one embodiment of the invention. The system includes client **410**, distributors **415-416**, traffic management devices **420-422**, and origin servers **440-442**.

Client **410** is coupled to distributor **415** over WAN/LAN **100**. Distributor **415** is coupled to distributor **416** through traffic management devices **420-422**. Distributor **415** also has a more direct connection to distributor **416**. Distributor **416** is coupled to origin servers **440-442**.

Client **410** is any device capable of connecting with WAN/LAN **100**. As such, client **410** is substantially similar to remote computer **140** in FIG. **1**.

Distributor **415** receives information in the form of packets. Each packet may convey a piece of information. A packet may be sent for handshaking, i.e., to establish a connection or to acknowledge receipt of data. A communication includes a group of related packets sent between two devices, such as client **410** and server **440**. For example, to request a Web page, client **410** may send some packets requesting a connection, e.g., handshaking packets, server **440** may respond with other handshaking packets. Then client **410** may send a packet requesting a particular Web page. Server **440** may respond by sending data packets associated with the Web page. Finally, client **410** may end a communication by sending some more handshaking packets which server **440** may respond to with other handshaking packets. In essence, a communication may be thought to include all packets needed or necessary for a transaction to occur. A communication or part of a communication may also be referred to as a flow or as a flow of packets.

As described above, a flow may include a bi-directional flow of packets. Bi-directional packet flows include packets sent from a client, such as client **410**, that are destined for a server, such as origin server **440**, and those packets sent from the server to the client. A flow of packets may also include related packet flows, such as a control packet flow and a data packet flow that may arise during a File Transfer Protocol (FTP) session, or the like. A flow of packets might further include IP fragments that may arise either at the original sender of the packets, or at any intermediate device along a communication path.

A packet may come from various senders including client **410**, traffic management devices **420-422**, distributor **416**, or origin servers **440-442**. The packet may include information such as a request, response, or configuration command. Generally, packets received by distributor **415** will be formatted according to TCP/IP, but they could also be formatted using another transport protocol, such as User Datagram Protocol (UDP), Internet Control Message Protocol (ICMP), NETh-eui, IPX/SPX, token ring, and the like. Upon receipt of a packet, a transcoder (not shown) associated with distributor **415** makes a determination as to where the packet should go. The transcoder may be logic circuitry or software executing on an electronic device, such as a processor, within distributor **415**, or it may execute or reside elsewhere.

In one embodiment, the transcoder includes traffic distribution engine **224** shown in FIG. **2**. The transcoder may access a database, a table, or other information to determine an action to perform upon receipt of a packet, or it may be "hard-wired" to perform a certain action depending pre-defined conditions. In some senses, the transcoder may be viewed as the "brains" of distributor **415** or as logic, which drives the operation of distributor **415**. In future references to distributors, the word transcoder may or may not be used. Furthermore, a distributor may be referred to as making decisions or determinations, but it should be understood in such references that a transcoder associated with the distributor is making the decisions or determinations and causing the distributor to take action appropriately.

A user on client **410** may request, for example, a Web page associated with a URL, such as http://www.uspto.gov. If this is a Web page that is serviced by origin servers **440-442**, distributor **415** forwards the request to one of traffic management devices **420-422**. A user on client **410** may request communication specifically with one of the traffic managers. In this case, distributor **415** forwards the request to the specified traffic manager. For example, the user may wish to configure the traffic management device, install new software, provide maintenance, or some other activity. The user may wish to configure distributor **415**. In this case, distributor **415** processes the communication itself. Distributor **415** may receive a response to a previous request from one of traffic management devices **420-422**. Distributor **415** may then forward this request to the recipient by sending it to WAN/LAN **100**. A user may send a message directed specifically at one of origin servers **440-442**. In this case, distributor **415** may send the message to distributor **416** for relaying the message to the specified server.

When requests for content come to distributor **415**, the distributor may be required to ensure that a request from the same source is sent through the same traffic management device. Distributor **415** (and **416**) may employ a variety of techniques to ensure that the request is from the same source is sent through the same traffic management device. For example, distributor **415** (and **416**) may employ such techniques as are described in U.S. patent application Ser. No. 10/119,433, filed Apr. 9, 2002, entitled "Method and System for Scaling Network Traffic Managers," which is hereby incorporated by reference.

Distributor **415** may extract information from the request to select the traffic management device. By dynamically employing the extracted information for each request, distributor **415** need not maintain state information about the connections between origin servers **440-442** and requesters such as client **410**. Distributor **415** may select the traffic management device and forward requests by performing actions described below in conjunction with FIG. **8**.

Sometimes, when distributor **415** receives a packet, it acts like a router or switch, forwarding the packet toward the intended recipient. For example, distributor **415** may receive

a request to connect with server **440**. Distributor **415** may forward this request to distributor **416** for forwarding to server **440**. Distributor **415** may receive a packet from distributor **416** or from traffic management devices **420-422** that is directed at a client, such as client **410**. In this case, distributor **415** forwards the packet to WAN/LAN **100** (or a router thereon). Alternatively, if client **410** is a device distributor **415** is more closely connected to, distributor **415** may send the message directly to client **410**.

Distributor **415** may use a different algorithm for forwarding messages directed towards traffic management devices **420-422** than for messages from traffic management devices **420-422**. For example, when messages are directed at traffic management devices **420-422**, distributor **415** may perform a hash on either a source or destination IP address and port number to determine to which traffic management device the message is to be sent. When distributor **415** receives a message from a traffic management device, however, it may forego applying a hash.

A hash is a function or algorithm that maps a set of input values to a set of output values. Typically, a hash is used when the set of input values has more elements than the set of output values. Some hashes when applied to a set of input values will map the input values approximately equally over the set of output values. Other hashes will map the input values disproportionately to a set of output values. For example, one traffic management device may be able to deal with twice as many packets as another traffic management device. A hash could be constructed to map input packets to the one traffic management device twice as often as mapping packets to the other traffic management device. Generally, a hash is deterministic. That is, the hash will produce the same output value whenever a particular input value is hashed on.

Traffic management devices **420-422** receive messages sent from distributors **415** and **416**. In some operations, traffic management devices **420-422** act like level 7 switches. That is, they may look at content associated with higher TCP/IP layers of the message, e.g. a request for a page such as http://www.uspto.gov/and information that identifies the user, such as a cookie, etc. They may store information in memory so that next time the requestor requests more information from http://www.uspto.gov/, each request is sent to the same server. They may do this, in part, to ensure that the user is connected to the server that the user previously connected to. This helps prevent the loss of transaction data, such as items in a shopping cart.

In addition, traffic management devices **420-422** may perform network address translation (NAT). That is, in a TCP/IP packet, they may change the source and/or destination field. This may be done for many reasons. One reason is that each traffic management device is configured to cause future communications to and from a server to flow through the traffic management device, so that the traffic management device may maintain state information about the connection. The traffic management device may need state information to gracefully close a connection if, for example, the server fails. In addition, the traffic management device may need state information to reroute a connection to another server if the server fails. Another reason the traffic management device may be configured to have all future communications flow through it is for security purposes.

When traffic management devices **420-422** perform a network address translation, a relationship between the source and destination field may change. For example, a source port number that was originally less than a destination port number, may now be greater. Such a change in the relationship between the port numbers may affect which traffic management device the reply is sent. This may result in a different traffic management device being selected to handle the reply than was selected to handle an initial client request. Therefore, traffic management devices **420-422** may need to perform actions such as those described below in conjunction with FIG. **9** to ensure that the same traffic management device handles packets in the same flow of packets.

Traffic management devices, such as traffic management devices **420-422**, are any devices that manage network traffic. Such devices include, for example, routers, proxies, firewalls, load balancers, devices that perform network address translation, any combination of the preceding devices, and the like. A traffic manager may, for example, control the flow of data packets delivered to and forwarded from an array of application servers, such as Web servers. A traffic manager may direct a request for a resource to a particular Web server based on network traffic, network topology, capacity of the server, content requested, and a host of other load balancing metrics. A traffic manager may receive data packets from and transmit data packets to the Internet, an intranet, or a local area network accessible through another network. A traffic manager may recognize packets that are part of the same communication, flow, and/or stream and may perform special processing on such packets, such as directing them to the same server so that state information is maintained. A traffic manager may support a wide variety of network applications such as Web browsing, email, telephony, streaming multimedia, and other traffic that is sent in packets.

A traffic management device may be implemented using one or more personal computers, POCKET PCs, wearable computers, multiprocessor systems, microprocessor-based or programmable consumer electronics, network PCs, cell phones, smart phones, pagers, walkie talkies, radio frequency (RF) devices, infrared (IR) devices, CBs, integrated devices combining one or more of the preceding devices, and the like. Such devices may be implemented solely in hardware or in hardware and software. For example, such devices may include some application specific integrated circuits (ASICs) coupled to one or more microprocessors. The ASICs may be used to provide a high-speed switch fabric while the microprocessors may perform higher layer processing of packets. An exemplary device that could be used as a traffic management device is server computer **200** of FIG. **2**, configured with appropriate software. A traffic management device may have multiple network interface units and each network interface unit may interface with one or more networks. It should be understood that traffic manager as it is used in this document means traffic management device.

Distributor **416** receives communications and forwards them to one or more of origin servers **440-442**, to distributor **415**, or to traffic management devices **420-422**. When forwarding messages to traffic management devices **420-422**, distributor **416** may hash on the IP address and port number of either the destination or source, depending on which is greater, so that the same traffic management device that sent a packet to a server receives the server's response. Distributor **416** may perform actions described below in conjunction with FIG. **8** to select the traffic management device. Distributor **416** may act as a switch or router in relaying messages to intended recipients. Although distributor **416** is shown as having one shared communications link (segment) going between it and origin servers **440-442**, it may have dedicated communications links to each of origin servers **440-442**.

Origin servers **440-442** may include one or more WWW servers, such as network device **200** of FIG. **2**, or other general-purpose servers. Origin servers **440-442** may serve content for more than one vendor. For example, a group of ven-

key and employing the hash key to select the first traffic manager, wherein a response packet to the received packet is forwarded to the first traffic manager; and

(iii) if the received packet is a packet of a second type, forwarding the received packet to a second traffic manager that is selected based in part on a second field in the received packet, wherein the response packet is forwarded to the second traffic manager and wherein the second field in the received packet is hashed to obtain another hash key that is used to select the second traffic manager.

16. The method of claim **15**, wherein receiving the packet further comprises determining the packet type of the received packet based in part on comparing source information with destination information in the received packet.

17. The method of claim **15**, wherein a packet of the first type further comprises a packet that includes a source port number that is greater than a destination port number.

18. The method of claim **15**, wherein a packet of the second type further comprises a packet that includes a destination port number that is greater than a source port number.

19. The method of claim **15**, wherein the first field in the received packet further comprises a source IP address and the source port number, and the second field in the received packet further comprises a destination IP address and the destination port number.

20. The method of claim **15**, further comprising, if the first field in the received packet equals the second field in the received packet, forwarding the received packet to a pre-determined traffic manager.

21. The method of claim **15**, further comprising, if source information in the received packet is to be translated, replacing a source port number in the received packet with another source port number, wherein the other source port number is greater than a destination port number associated with the received packet.

22. The method of claim **21**, wherein the other source port number is selected from a pre-computed self-source port table.

23. A system for routing a packet over a network, comprising:

(a) a plurality of servers;

(b) a plurality of traffic managers arranged to direct the packet to at least one of the plurality of servers; and

(c) a distributor, coupled to the plurality of traffic managers, that is arranged to perform actions, including:

(i) if the received packet is a first packet type, forwarding the received packet to a first traffic manager in the plurality of traffic managers that is selected using in part a first field in the received packet by hashing the first field to obtain a hash key used to select the first traffic manager, wherein a response packet to the received packet is forwarded to the first traffic manager; and

(ii) if the received packet is a second packet type, forwarding the received packet to a second traffic manager in the plurality of traffic managers that is selected using in part a second field in the received packet by in part hashing the second field to obtain another hash key used to select the second traffic manager, wherein the response packet is forwarded to the second traffic manager.

24. The system of claim **23**, wherein the first packet type includes a source port number in the received packet that is greater than a destination port number in the received packet, and a second packet type includes a destination port number in the received packet that is greater than a source port number in the received packet.

25. The system of claim **23**, wherein using in part the first field further comprises using a source IP address and a source port number in the received packet.

26. The system of claim **24**, wherein using the part the second field further comprises using a destination IP address and a destination port number in the received packet.

27. The system of claim **23**, wherein in part hashing the first field further comprises:

hashing a source IP address and a source port number in the received packet to obtain the hash key; and

employing the hash key to select the first traffic manager to which the received packet is forwarded.

28. The system of claim **27**, wherein hashing the source IP address includes employing a hash function that is configured to load balance the plurality of traffic managers.

29. The system of claim **23**, wherein the distributor is arranged to perform actions, further comprising, if the received packet is associated with a pre-determined group characteristic, selecting the traffic manager and the other traffic manager from a plurality of traffic managers that are partitioned into groups of traffic managers based in part on the pre-determined group characteristic.

* * * * *